Gender differences in the choice of field of study and the relevance of income information: Insights from a field experiment

Finger, Claudia; Solga, Heike; Ehlert, Martin; Rusconi, Alessandra

Finger, Claudia; Solga, Heike; Ehlert, Martin; Rusconi, Alessandra

Article — Accepted Manuscript (Postprint)

Gender differences in the choice of field of study and the relevance of income information. Insights from a field experiment

Research in Social Stratification and Mobility

Provided in Cooperation with:
WZB Berlin Social Science Center

Suggested Citation: Finger, Claudia; Solga, Heike; Ehlert, Martin; Rusconi, Alessandra (2019) : Gender differences in the choice of field of study and the relevance of income information. Insights from a field experiment, Research in Social Stratification and Mobility, ISSN 1878-5654, Elsevier, Amsterdam, Vol. Articles in press, Iss. [Article No.] 100457, pp. 1-17, http://dx.doi.org/10.1016/j.rssm.2019.100457

This Version is available at:
http://hdl.handle.net/10419/209685

Accepted for publication in *Research in Social Stratification and Mobility*,
https://doi.org/10.1016/j.rssm.2019.100457 (available online 29 November 2019)

Gender differences in the choice of field of study and the relevance of income information. Insights from a field experiment

Claudia Finger *,[a], Heike Solga [a, b], Martin Ehlert [a], Alessandra Rusconi [a]

[a] WZB Berlin Social Science Center (WZB, Reichpietschufer 50, 10785 Berlin, Germany)

[b] Freie Universität Berlin (Institute of Sociology, Garystr. 55, 14195 Berlin, Germany)

Abstract

Research consistently reports pronounced earnings differences between men and women, even among the highly educated. This article investigates whether students' responsiveness to information on income returns relates to gender differences in major choices, which might contribute to the persistent gender wage gap. We use field-experimental panel data on students in Berlin (Germany), starting one year before high school graduation. Our intervention comprised information on major-specific returns to college and was provided to students in randomly selected schools. By comparing the major-specific application decisions of "treated" and "untreated" high school seniors, we examine whether, and why, male and female students respond differently to this information. As potential mechanisms behind a gender-specific treatment effect, we analyze the role of gender stereotypes and roles associated with certain job attributes. We find that providing income information on college majors only influences the major choices of male (not female) students with college intention: treated male students on average applied to majors associated with higher mean income. Further analyses suggest that this gender difference in the treatment effect cannot be explained by differential distributions or effects of preferred job attributes.

Keywords: Gender inequality; College major choice; Gender roles and stereotypes; Monetary returns; Information; Field experiment

1 Introduction

The gender wage gap is persistent and pronounced in advanced societies (OECD, 2018). One important reason is that men choose fields of study, or college majors, that lead to more lucrative occupations in terms of income than women (e.g., Bobbitt-Zeher, 2007; Leuze & Strauß, 2009). Thus, improving our knowledge about the mechanisms behind the gendered patterns of major choices has been identified as highly relevant to combating gender inequality in the labor market (e.g., Barone, 2011; Charles & Bradley, 2002, 2009; Jonsson, 1999; Legewie & DiPrete, 2014; Lörz, Schindler, & Walter, 2011; Mann & DiPrete, 2013; Morgan, Gelbgiser, & Weeden, 2013; Ochsenfeld, 2016; Zafar, 2013). In this paper, we contribute to this literature by exploring students' responsiveness to information on major-specific wages for gender differences in major choices. To this end, we use data from a field experiment in which we provided information about major-specific wages to a randomly selected group of high school students.

Research on gender differences in major choices (e.g., Barone, 2011; Ochsenfeld, 2016) draws on two theoretical perspectives: cultural explanations and rational choice explanations. The former state that boys and girls internalize gender stereotypes and roles during their socialization, resulting in different interests, course work patterns, (subjectively perceived) abilities, and life goals that, in turn, lead to gender differences in major choices (e.g., Charles & Bradley, 2002, 2009). According to rational choice explanations, by contrast, gendered major choices are due to gender-specific perceptions of social costs and success probabilities (Jonsson, 1999).[1] A further argument is that male and female students evaluate the benefits of certain majors and related occupations based on different criteria, such as wage penalty expectations due to career interruptions (Polachek, 1981), gender-typical interests (e.g., orientation towards objects or people), or life plans (breadwinner vs. homemaker/caregiver) (e.g., Lörz et al., 2011).

Empirically, various studies agree that gender stereotypes and roles seem to lead to "gender-appropriate" decisions for, or against, certain majors (e.g., Charles & Bradley, 2009; Ceci, Williams, & Barnett, 2009; Jonsson, 1999; Lörz et al., 2011; Ochsenfeld, 2016). What is still unclear, however, is whether they result in "gender-appropriate" major choices because

[1] Social costs include belonging to the minority and fear of discrimination (Jonsson, 1999, p. 394). Success probabilities differ because of "sex-specific competitive advantages" in different subjects, that is, girls and boys "prefer to specialize […] in the subject for which they obtain their highest school marks," which eventually channel them into different fields of study; Jonsson, 1999, p. 395).

of the aforementioned cultural or rational choice explanations. If cultural mechanisms are the driving source, progress towards ungendered major choices seems to be quite unlikely (Barone, 2011; Charles & Bradley, 2002; Lörz et al., 2011).

Most scholars acknowledge that educational decisions are taken under uncertainty and incomplete or incorrect information so that rational decision-making might be biased. Research has shown that students often have incorrect expectations about income returns to higher education (HE) (e.g., Hastings, Neilson, & Zimmermann, 2015; Oreopoulos & Dunn, 2013). Correspondingly, experimental research shows that providing detailed information on various, often monetary outcomes "nudges" students to make better informed and thus partly different educational choices (e.g., Davies, Davies, & Qiu, 2017; Domina, 2009; Hastings et al., 2015; McGuigan, McNally, & Wyness, 2016). Even though evidence is inconsistent (see Herbaut & Geven, 2019 for a review), some studies reveal that additional financial information on HE especially increases the college intentions, applications, or attendance of students from socially disadvantaged families (or from broader disadvantaged contexts) (e.g., Ehlert, Finger, Rusconi, & Solga, 2017; Loyalka, Song, Wei, Zhong, & Rozelle, 2013; Oreopoulos & Dunn, 2013; Peter & Zambre, 2017). We know little, however, about whether and why the impact of detailed information on major-specific earnings differs between male and female students. To our knowledge, only two papers addressed this question so far: Barone, Schizzerotto, Assirelli, and Abbiati (2019) for Italy and Kerr, Pekkarinen, Sarvimäki, and Uusitalo (2014) for Finland.

In this article, we contribute to the literature in several ways. First, we study the impact of an information treatment about major-specific returns on major choices of young men and women in Germany, adding to the cumulative knowledge on this topic by providing evidence from a different institutional context. Second, we measure major choices and their financial lucrativeness in a more direct and fine-grained way than previous studies: as field-specific average wages. Third, we pay particular attention to the interplay of gender stereotypes and roles and the role of income information for college major choices. We use field-experimental panel data on students in high schools with a college-preparatory track in Berlin, Germany. In our field experiment, we provided information on (major-specific) returns to HE to students in randomly selected schools one year before they obtained their HE entrance certificate and followed them until one and a half years after leaving general schooling.

2. Previous research and theoretical considerations

2.1 Income information and major choice: findings from experimental studies

Economic research, focusing on monetary returns, shows that major choices are associated with income expectations (Berger, 1988; Hastings et al., 2015; Hastings, Neilson, Ramirez, & Zimmerman, 2016; Huntington-Klein, 2016; Montmarquette, Cannings, & Mahseredjian, 2002; Zafar, 2013). Yet such expectations better explain the major choices of male students—indicating that income considerations are more relevant to young men than women (Montmarquette et al., 2002; Zafar, 2013). At the same time, major-specific income expectations are strongly biased—for both male and female students—meaning that major choices are often based on wrong expectations (Betts, 1996; Hastings et al., 2015; Wiswall & Zafar, 2015a, b).

A growing number of researchers use field-experimental designs to study whether reducing such information biases leads students to make better-informed decisions. They mostly provide information on the costs of, and returns to, different kinds of education and test its effect on HE intentions and decisions (e.g., Barone, Schizzerotto, Abbiati, & Argentin, 2017; McGuigan et al., 2016; Oreopoulos & Dunn, 2013). Although some studies also look at the impact of information on major choices (Barone et al., 2019; Hastings et al., 2015; Kerr et al., 2015; Wiswall & Zafar, 2015a, b; for A-level choices see Davies et al., 2017), most of them do not differentiate their analysis by gender.

One study that did was conducted by Barone and colleagues (2019), who investigated whether providing information to Italian high school seniors has a gender-specific impact on their major choices. The authors hypothesized that the information treatment might have a stronger effect on girls than on boys in terms of choosing more rewarding majors, because boys' preferences already overlap more strongly with high(er) rewarding majors/occupations, meaning boys are more likely to choose such fields irrespective of information biases. Thus, assuming an information deficit among students, female students should benefit more from such financial information, which might stipulate rational evaluations and redirect them towards more rewarding fields (Barone et al., 2019, p. 360).

The intervention was quite extensive, lasting five hours, spread over three meetings. Besides information on costs and success probabilities, it provided various information on employment prospects for different fields of study (such as first job search duration, earnings, and risk of vertical and horizontal mismatch), broken down into strong fields (engineering, computing, and medicine), intermediate fields, and weak fields (humanities and social sciences (Barone et al., 2019, p. 362). Their results show female students to be neither less informed than men before the treatment nor to revise their expectations about returns to education more strongly than men (Barone et al., 2019, pp. 370-371). However, they find that the information treatment did reduce the rates of girls (but not of boys) signing up for entrance tests and enrolling in so-called weak fields of study (Barone et al., 2019, p. 367). Yet neither girls nor boys were more likely to apply to, or enroll in, majors categorized as "strong." However, it remains unclear whether these results support rational-choice explanations, because they indicate that "treated" girls did change their major choice but still opted for fields that fit better to their (gendered) preferences (e.g., teaching or psychology rather than hard sciences).

Likewise, the study by Kerr and colleagues (2014, 2015) does not point towards such an explanation. For Finland, they find that earnings information does not encourage students to enroll in more rewarding majors or sign up for entrance tests for such fields; if at all, only "treated" boys from less-educated neighborhoods "apply more to fields with better labor market prospects" (Kerr et al., 2014, p. 18). In contrast to the Italian study, they used application profiles based on fine-grained actual income differences between college majors and probabilities of success in entrance exams as dependent variable. Moreover, their intervention only took 20-25min. The main goal was thus to make students more aware of their information deficits—successfully so, according to the findings (Kerr et al., 2015)—and consequently encourage them to do further research. In addition, access to university is restricted in Finland, meaning students always have to register for entrance exams first.

These contradictory results may result from institutional context or design differences (e.g., different dependent variables, different intensity of the intervention). However, both studies have two limitations: First, they do not differentiate between students with and without college intentions. According to Hanson (1994, p. 159), students with college intentions are those who *expect* to enroll in college, based on everything they know. To take students' college intentions into account might be important, however, because the interventions were conducted late in their high school career. At this stage, educational plans might be more consolidated than earlier (e.g., due to previous course-taking patterns, see Jonsson (1999)), meaning that opposing information might be less influential for educational decisions. Correspondingly, we know from our experimental study that the information intervention only increased the application and enrollment rates of students with college intentions (Ehlert, Finger et al., 2017; Peter, Rusconi, Solga, Spieß, & Zambre, 2016). Thus, including all students in the analyses might underestimate the impact of financial information on those with college intentions. Second, both studies focused on rational choice explanations and not on gender stereotypes and roles (as part of cultural explanations) as potential mechanisms underlying gender differences in the responsiveness to income information.

With our analysis, we extend this previous research in several respects. First, we explore the impact of an information intervention on wage differences between college majors for male and female students' HE applications (our operationalization for "major choice"). In contrast to most previous research, we deploy a fine-grained measure of college majors, differentiated by their actual income returns as outcome variable. This measure captures income-related changes in major choices within broader major categories that would otherwise be hidden and, at the same time, does not overestimate small income differences, even if the respective majors belong to very different categories (see also Kerr et al., 2015). Second, as our design (see Section 4) resembles that of the Finnish study, we add cumulative evidence on short-duration information interventions (or awareness-raising interventions), and we explore the generalizability of the Finnish finding (treatment effect especially for boys from less-educated neighborhoods). Third, as discussed below, we study the interplay between providing unbiased information (or information deficits) and gender stereotypes and roles that we approximate by differences in job attribute preferences.

2.2 Job attribute preferences and major choice

In the following, we consider the interplay between students' job attribute preferences, information on monetary returns, and gender differences in major choices. Job attribute preferences can be understood as "the extent to which people desire a variety of specific qualities and outcomes from their paid work" (Konrad, Ritchie, Lieb, & Corrigall, 2000, p. 593). Students might, for instance, prefer jobs that lead to a high income, provide opportunities for promotion, or allow for enough family time. Preferences for certain job attributes are associated with occupational and thus major choices: Students who value a high income, for instance, are on average more likely to opt for fields that lead to a higher income than students for whom money is less important (Daymont & Andrisani, 1984; Hastings et al., 2016; Ochsenfeld, 2016).

At the same time, and in line with cultural accounts, some job attributes relate to deeply rooted gender stereotypes and roles (Konrad et al., 2000). *Gender stereotypes* are the result of societal norms that are continuously transferred during socialization by parents' and teachers' expectations and ascriptions of male and female character traits and talents. Stereotypes linked to masculinity are, for instance, dominance, physical strengths, and the capacity for analytical reasoning. Feminine stereotypes refer, by contrast, to emphatic or pro-social behavior. Young women are often interested in social or altruistic tasks and thus prefer jobs that involve social interaction and the opportunity to help others (Barone, 2011; Bradley, 2000; Konrad et al., 2000; Ochsenfeld, 2016). This might be one reason for women's overrepresentation in majors such as humanities, social sciences, social work, teaching, and some health-related fields—fields that mostly do not lead to high-income jobs.

Gender roles might be another reason for different job attribute preferences. The common role of men as breadwinners and women as homemakers and caregivers contributes to men's preferences for well-paid jobs versus women's preferences for jobs that allow reconciling work and family duties (Hakim, 2002; Konrad et al., 2000; Lörz et al., 2011). However, due to an increasing labor-market integration of women, earned income should have become more important to women as well, while performing family duties should have become more valuable to men (Konrad et al., 2000). Nowadays, younger generations aim for an egalitarian division of labor—although, after the birth of the first child, care and household duties often become gendered again (e.g., Jansen & Liefbroer, 2006). Nonetheless, according to Hakim (2002), today's women often want to reconcile work and family. Thus, even if women consider a high income an important job attribute, further (and partly contradictory) attributes that emphasize their caregiving role more often come into play, potentially relativizing monetary goals.

In line with (cultural) theoretical expectations, previous research shows that male and female students differ in their job attribute preferences and related life goals (e.g., Bobbit-Zeher, 2007; Lörz et al., 2011; Morgan et al., 2013). Differences are stronger with regard to *stereotypical* job attributes such as having "contact to people" or "the opportunity to help others" and somewhat less pronounced when they refer to *gender roles* of "achieving a high income" or "having enough family time" (Konrad et al., 2000). Correspondingly, research shows that stereotypical vocational interests (e.g., for "working with machines" or "caring for people in need") explain a substantial part of the gender gap in major choices (Ochsenfeld, 2016), whereas broader (rather gender-role-related) work values or life goals such as "making money” or “having children” seem to be less important (Mann and DiPrete, 2013; Morgan et al., 2013; but see Lörz et al., 2011).

These different findings might result from differences in the gendering of the stereotypical versus role-specific job attribute preferences reported above. They might, however, also result from differences in the strength of the link between interests or life goals and college major categories (Ochsenfeld, 2016, p. 126). Vocational interests are closely linked to common categorizations of majors (e.g., “working with machines” connects to engineering or “caring for people in need” to social work or medicine). Broader life goals, however, fit less neatly into the common categorizations, as such goals can be realized through different majors (e.g., “making money” can be achieved via engineering and medicine).

2.3 Gendered preferences—gendered information processing? Our hypotheses

Next, we discuss the conditions under which we expect information on monetary returns to influence major choices and why gender roles and stereotypes might contribute to gender differences in this respect. First, we argue that the potential impact of monetary information varies depending on whether that information matches male and female students’ preferences for certain job attributes. Second, the relevance of such a match might depend on whether these job attribute preferences are in line with common gender roles and stereotypes (i.e., their gender conformity). We focus on job attribute preferences as an indication of a cultural explanation. This does not mean, however, that other sources, like admission barriers, course work in school, or grades, might not generate gender differences in major choice and in the effect of monetary information.

Attention being a scarce resource (DellaVigna, 2009), individuals mainly process information that is relevant to them because it matches their preferences (e.g., for certain job attributes) and tend to ignore information that is not (Hastings et al., 2016; McGuigan et al., 2016). This does not necessarily mean that students actively search for income information even if they prefer to get a job with a high income later. In fact, they often do not engage in information gathering and thus have strongly biased return expectations (e.g., Hastings et al., 2016; McGuigan et al., 2016). However, they might still be responsive to systematic and reliable information when provided without further costs (as in our information treatment). Why might this lead to gender differences in the likelihood to integrate income information into students’ major choice?

As discussed in the previous section, due to gender roles and stereotypes that are transmitted and internalized during socialization, women are more likely than men to express social, altruistic, and family-related preferences (rather than income preferences). Hence, female students' preferences might more often lead girls to interpret monetary information as not relevant and to ignore such information in their major choices. Male students, in contrast, are more likely to have income-related job attribute preferences. As a consequence, and in contrast to rational choice-oriented expectation (see Section 2.1), not female but male students might be more attentive to financial information, and more likely to integrate this information into their decision. We therefore expect *the provision of information on monetary returns to have a stronger positive effect for male than for female students regarding their choice of financially more lucrative majors (H1).*

As outlined in the previous paragraph and following a distributional argument, job attribute preferences could *mediate* the (potentially) gender-specific treatment effect formulated in H1. This could be the case if preferences for the job attribute "high income," for instance, are distributed differently among male and female students (i.e., if more male than female students state a preference for jobs leading to a high income) *and* if this preference increases the treatment effect. Accordingly, we expect that the *gender difference in favor of male students in the effect of providing income information (as stated in H1) becomes smaller when adjusted for gender differences in the (pre-treatment) distribution of job attribute preferences (H2, mediation).*

This hypothesis implies that job attribute preferences have the same meaning regardless of students' gender. As a result of gender socialization, the meaningfulness of such preferences for male and female students and thus their impact on the treatment effect could, however, also vary with their gender conformity. If this is the case, job attribute preferences might *moderate* the gender-specific impact of monetary information, that is, our information treatment might have a different impact on men and women with the *same* job attribute preferences.

On the one hand, and in accordance with cultural explanations, job attribute preferences might be especially meaningful when they are in line with common gender roles and stereotypes. Despite steady changes, preferring a high income is still more strongly aligned with common gender role expectations for men than for women (Konrad et al., 2000). Hence, preferring a high income might be particularly salient for men and strongly guide their decisions. If such a gender-conforming preference meets matching information, both could reinforce each other: In this situation, male students might be particularly attentive and responsive, which might further increase the effectiveness of information on major-specific income returns. Women, in contrast, most often try to reconcile family and work duties (Hakim, 2002). Thus, compared to their male peers, income preferences stated by women might be less salient than competing preferences that are in line with female roles and stereotypes like "spending time with family" or "having contact with people." Consequently, even if women state preferences for a high income and even if they get compatible information, this might less often increase their attentiveness and responsiveness to such information. These considerations result in the following hypothesis: *The gender difference in favor of male students in the effect of providing income information (as stated in H1) is stronger within the group of students with income-related job preferences (compared to students without such preferences) (H3a, moderation).*

Preferences for job attributes associated with female roles and stereotypes might be more salient for young women than men. If this is the foundation that meets (incompatible) income information, this information is likely to be interpreted as irrelevant and thus ignored or not integrated into the decision-making process. Thus, having gender-conforming job attribute preferences could *decrease* the impact of income information on women's major choices. For men, the same preferences are in conflict with common role expectations and might therefore be less salient. Consequently, such job attribute preferences, even if stated, might rarely affect men's responsiveness to income information. We therefore also expect the *gender difference in favor of male students in the effect of providing income information (as stated in H1) to be stronger within the group of students with job attribute preferences that comply with female stereotypes and roles (as compared to students without such preferences) (H4a, moderation).*

On the other hand, job attribute preferences might be especially meaningful when they are in conflict with common gender roles and stereotypes. According to psychological research, stating and complying with gendered stereotypes and roles does not necessarily mean that they are part of a person's identity and self-concept (Konrad et al., 2000). In this case, conforming to gendered expectations could only be the "default" to avoid social sanctions, and gender-conforming responses in a survey questionnaire might just be "empty words" to meet societal expectations (Konrad et al., 2000). Non-gender-conforming statements on job attribute preferences, by contrast, might be truly meaningful, because they are enforced against societal conventions and might therefore reinforce the impact of compatible information. If this were the case, we might observe the opposite, namely that stating preferences for high-income jobs would especially enhance the effect of providing information among women, whereas stating preferences for job attributes that are in line with female roles and stereotypes might especially decrease the effect of income information among men. We thus expect *the gender difference in favor of male students in the effect of providing income information (as stated in H1) to be weaker within both the group of students with income-related job preferences (H3b, moderation) and the group of students with job attribute preferences that comply with female roles and stereotypes (H4b, moderation).*

Support for H3a and H4a would suggest the strong relevance of gendered socialization, which leads to internalized "gender-appropriate" roles and stereotypes and might contribute to (in)attentiveness to (in)compatible information. Support for H3b and H4b, by contrast, would indicate that preferences corresponding to gender roles and stereotypes—despite being routinely reproduced in social encounters—less often lead to an active engagement with information than preferences that do not correspond to gender norms and thus arguably develop in a rather proactive way.

3 Institutional context

In the following, we describe the institutional conditions under which German upper secondary school graduates choose college majors. The figures reported refer to 2014—the year in which the students in our study gained HE eligibility.

The main pathway to HE eligibility in Germany is to obtain a so-called university entrance certificate—the *(Fach-)Abitur*—from an upper secondary school (including the traditional "Gymnasium," comprehensive schools with a "Gymnasium" track, and vocational "Gymnasium"). Because of the highly stratified German (secondary) school system, only 53 percent of all school leavers obtained such a certificate in 2014 (National Education Report, 2018, Tab. F2-1A).

At the beginning of their penultimate school year, students choose their course profile, consisting of advanced and basic courses. The students' course profiles do not formally restrict their application for certain majors, but they might influence their major choices. Students who plan to attend HE directly after high school apply to college programs in summer, shortly after their graduation. Those students who intend to start an apprenticeship instead need to apply during their last school year (between December and May). Career guidance activities during upper secondary education differ quite strongly between the German states. Our experimental intervention was conducted in 2013 in Berlin (see Section 4.2). At that time, most upper-secondary schools were not required to offer systematic career activities to all students in Berlin.[2] They were allowed to offer a supplementary course on "Studying and Occupational Career," which students could choose voluntarily. Moreover, about 20 percent of Gymnasium students and about 30 percent of comprehensive school students participated in the voluntary state program "In-depth Vocational Orientation" (Böhm & Pampel, 2014, p. 9). Furthermore, the German Federal Employment Agency provides information material and offers occupational counseling for individuals and school classes. Classes, for instance, visit so-called "Job Information Centers," where students can autonomously retrieve information on occupations (Saniter, Schnitzlein, & Siedler, 2019). However, these information-gathering activities are focused on the requirements and contents of occupations rather than on income returns to college majors. Moreover, they usually take place at least two years before the end of upper secondary school and thus before our treatment (see Section 4.4).

Despite strong selection into upper secondary schools, only around 70 percent of college-eligible students actually enrolled in HE programs: most of them either directly after graduation (45 %) or after one gap year (23 %) (National Education Report, 2018, Tab. F2-6web, F2-21web). One explanation for the low enrollment rates is the attractiveness of the German apprenticeship system, which diverts certain students from HE (Mayer, Müller, & Pollak, 2007; Powell & Solga, 2011). In terms of gender, more female (58 %) than male (48 %) school leavers obtained the HE entrance certificate, whereas, among them, more men (77 %) than women (69 %) eventually enrolled in HE. Overall, this leads to gender parity in German HE (National Education Report, 2018, Tab. F2-1A, F2-2A, F2-6web).

Unlike the school system, the German HE system is much less stratified. It can be categorized as binary, with two main institutional types: traditional full (or research) universities and universities of applied sciences with a limited range of fields of study (Mayer et al., 2007). Compared to other countries like the US, UK, or France, differences between universities with regard to institutional prestige and quality are (still) rather small. The German HE system is, however, strongly differentiated in horizontal terms with regard to fields of study. Undergraduate programs are rather narrowly defined. Hence, the far-reaching

[2] Systematic career guidance activities started in Berlin in 2015/16 (see http://www.psw-berlin.de/fileadmin/content/Downloads/landeskonzept/landeskonzept.pdf; accessed 2019/07/17).

decision for a specific major, and in many cases for a related occupation, has to be made during the application stage, or at the latest when students enroll. The majors differ in several dimensions: The formal time to degree, and thus direct and indirect study costs, for instance, ranges from three years (majors that lead to a bachelor's degree) to six years (e.g., law or medical degrees). Admission barriers also vary strongly by major. If the demand for a certain program exceeds its capacities, HE institutions are allowed to restrict admission. When such a *numerus clausus* (NC) applies, applicants are ranked according to certain criteria (mainly their average school grade) and are admitted until the pre-defined number. In 2013, the overall share of NC programs was around 50 percent (ranging from 40 % in language and cultural sciences to 100 % in medical programs). Due to these capacity constraints, major choices are not always "free" choices of individual students but "strategic" choices based on students' preferences *and* their expectations of success regarding college admissions and completion—a fact that might limit the impact of interventions aimed at altering individual decisions.

The pronounced gender difference in major choices, mentioned in the previous sections, also applies to Germany: the share of female students ranges from around 20 percent in engineering to 80 percent in education (National Education Report, 2018, Tab. F2-12web). At the same time, graduating in certain college majors is much more strongly associated with income returns than graduating from specific institutional types or single HE institutions (Spangenberg et al., 2012). The gendering of majors can therefore be expected to contribute to the gender wage gap.

Given the pronounced horizontal differentiation of the German HE system, the strong association between majors and income returns, and the marked gender differences in major choice, Germany is an interesting case to illuminate the impact of providing monetary information on gender differences in students' major choices.

4 Data and methods

To test our hypotheses, we use data from the "Best Up" study (Berliner-Studienberechtigten-Panel), which combines a panel survey of secondary school students with a randomized information treatment. In this section, we first describe the study's sample and experimental design (for details see Ehlert, Peter et al., 2017), then introduce the variables, and finally detail the empirical methods.

4.1 Sampling procedure and context

In the Best Up study, we collected data from upper secondary school students in 27 Berlin schools that lead to a HE entrance certificate (*Abitur* or *Fachabitur*) and followed them for five years.

To obtain the Best Up sample, we stratified existing schools using (1) school type; (2) share of adult population (>24 years) with low education (ISCED 0-2) per district (ranging from 7 % to 30 % in Berlin); (3) cohort size; (4) share of students with a migration background; and (5) share of female students as stratifying variables. Best Up focuses on students from lower-educated families. Therefore, the sampling focused on strata with an above-average share of lower educated adults (17 % and higher). Since residential segregation in Berlin is low in international comparison, these quarters are not heavily deprived. Nevertheless, families are on average poorer than in more advantaged districts of Berlin. It is thus important to note that we cannot extend the findings of our study to students from wealthier neighborhoods. Our sample is neither representative of Germany nor of Berlin. Thus—as with most field-experimental studies—our findings do not refer to a well-defined population, and significance tests are mainly used to identify the precision of our estimates. Referring our findings to a well-defined population is, however, not the main goal of the current study; we rather want to test theoretically derived mechanisms behind heterogeneous treatment effects.

Restricting our study to Berlin has the advantage of ruling out confounding influences caused by pronounced differences between (and within) the German states in terms of, for instance, existing school types, selectivity of and coverage with HE institutions and college majors (e.g., Helbig & Nikolai, 2015). It is, for instance, empirically well-established that physical proximity to universities and majors affects participation rates (e.g., Denzler & Wolter, 2010; Spieß & Wrohlich, 2010). Berlin has four research universities and 27 universities of applied sciences offering the whole range of majors. Overall, focusing on Berlin enables us to exclude the possible impact of long distances to HE institutions and majors on major-specific application decisions.

4.2 Survey

Within the 27 schools, we collected data from all students at the end of their penultimate school year—grade 11 or 12, depending on the school type—as in May/June 2013 (survey mode: paper and pencil, sample size: 1,578, response rate: 60 %). We re-contacted the students four times for follow-up online surveys: at the beginning of the final school year, shortly after high school graduation, at the beginning of the (potentially) second semester, and during the (potentially) third semester. Taking the first wave as a reference point, the response rate of the following online surveys lies between 70 percent (wave 2) and 62 percent (wave 5). Importantly, characteristics of panel dropouts do not vary significantly between treatment group (TG) and control group (CG) (see Ehlert, Peter et al., 2017).

In our analyses, we use the pre-treatment survey (wave 1) to measure students' job attribute preferences and initial major intentions and waves 3-5 to measure the major applied to up until one year after high school graduation, that is, after one gap year. Although we cannot include later applications, for instance those following an apprenticeship, we do cover a large share, as most students who eventually enroll in HE do so either directly after high school or one year later (see Section 3).

4.3 Sample definition

To test our hypotheses about gender-specific major choices in HE, we restrict our sample to students who intended to enroll in college in the first wave[3] and named the major they intended to study. We focus on students with a college intention in the first wave (77 % of male and 74 % of female students), because we assume that students only pay attention to information if they think it is relevant to them (Hastings et al., 2016; McGuigan et al., 2016). This may even be amplified by different application timelines between those who intend to go to college and those who intend to start an apprenticeship, because the latter needed to engage with and apply to apprenticeship positions shortly after the intervention took place (see Section 3). We would therefore also not expect to find an impact on major choices for those

[3] We measure college intentions by the following question from the German National Education Panel Study (NEPS, A49_T_Panel_2012©NEPS; see Stocké, Blossfeld, Hönig, & Sixt, 2011). "Based on everything you know now: What type of education will you probably pursue after leaving school? If you're planning to do a voluntary social year, an internship, or the like when you finish school, please choose the type of education you will probably pursue afterwards."

who did not intend to go to college. In line with this, previous analyses have shown that our information intervention only influenced students who stated such intentions in the first wave, that is, it "stabilized" college intentions stated at the end of the penultimate school year, but did not change the application decision of those who intended to start an apprenticeship (Ehlert, Finger et al., 2017; Peter et al., 2016). Nevertheless, we conducted a robustness check including these students; the results did not change (see Section 5.3).

As we are interested in major choices, we additionally restricted the sample to those respondents for whom we know that they applied to college and the major to which they applied.[4] These restrictions left us with a sample size of 557 respondents. By excluding those with missing information on the variables used for the analyses, our final analysis sample consisted of 510 students. Table A1 in the appendix documents how sample selection and panel attrition influenced the composition of the sample compared to the initial sample.[5] Overall, the analysis sample does not deviate substantively from the excluded cases except for factors known to be related to college attendance, such as gender and grades. Most importantly, the treatment status does not systematically vary between included and excluded cases (31.6 % and 29.4 %, respectively; see Table A1 in the appendix).

4.4 Experimental design and treatment

We assigned the information treatment to nine randomly selected schools (three per upper secondary school type, see Section 3). The randomization was stratified by school type, the average educational level of the neighborhood, cohort size, share of students with a migration background, and share of female students. Due to communication problems with the school

[4] As a consequence of this restriction, we also exclude students who entered VET even if they initially planned to go to university (N: 113). Adding the choice between VET and HE would add complications to the analysis. Since VET is shorter and students earn a wage during the course, we would have to take into account differences in (opportunity) costs in addition to the differences in returns that we are interested in.

[5] We decided against imputing because the number of cases lost due to item non-response is very small (about 4 % of the initial sample). Most of the cases from the initial sample are lost due to unit non-response (panel attrition; see Table A2 in the appendix) and due to missing values on college majors (the dependent variable). For both reasons, imputation is not recommended (see Von Hippel, 2007; Young & Johnson, 2015). Our balancing strategy addresses observed selective attrition between the TG and CG (see Section 5.6).

staff, however, we could only carry out the treatment in eight schools. The remaining 19 schools in our sample served as the control group. In our analysis sample, 161 students belong to the TG (66 male, 95 female) and 349 students to the CG (137 male, 212 female). We will discuss the covariate balance between TG and CG in Section 4.6 below.

The treatment took place in the classroom, directly after the first survey, that is, rather late in the school career (e.g., after the selection of advanced courses, see Section 3). It consisted of a 20-min presentation on the returns to, costs of, and ways to finance HE and vocational education and training (VET). It was given by researchers, who are perceived as credible authorities, thus enhancing students' confidence in the accuracy of the information provided (see Kolsto, 2001; Morgan, 2010). For the present study, it is important that the presenters provided, first, information on the income returns to a HE and VET degree (on average and over the life course). Second, they explained average income differences between men and women with both a HE and a VET degree, also mentioning the main reasons for women's lower income (more often employed in occupations with lower income, underrepresentation in leadership positions, and more frequent interruptions for childcare). Third and most importantly, they presented income returns of several majors and apprenticeship occupations based on survey data (average monthly net income of full-time employees), pointing out that wages vary considerably among college graduates, depending on their major. Figure 1 depicts three slides used to provide this information.

[**Figure 1**]

The presentation was followed by a three-minute film repeating the main messages: similar costs for HE and VET, advantages of HE particularly with regard to monetary returns (over the life course and in most majors), and ways to finance HE. Both the TG and CG received a one-page flyer with some general information on college attendance, common post-school opportunities, and a short list of websites with further information on financial aid options as a baseline treatment to level out differences in knowledge about where to find relevant information. We thus compare the impact of our face-to-face information treatment to a control treatment that only involved the written information flyer. By providing this basic treatment, we also reduce a potential ethical problem of our field experiment by not excluding the CG from useful information.

We do not expect that our short treatment *in itself* affects students' major choice. We rather believe that we made the treated students aware that they were lacking important and correct information (Morgan, 2010) and induced them more often (than the untreated students) to start searching for more information (available online).[6]

4.5 Variables

The main *dependent* variable refers to students' major choice. It measures the average income returns that students can expect after graduating in a chosen major. We focus on *major application decisions*, because applications do not confound individual choices and institutional admission decisions (which are not the target of our intervention). This is not to say that applications are a pure measure of individual decisions (or "pure" preferences), as they also capture strategic behavior, anticipated barriers, or discrimination that students include in their application decisions (e.g., Boliver, 2013). Yet they are more closely linked to individual preferences than enrollment. In waves 3 and 5, we included questions on the majors that participants applied to either directly after high school graduation or one year later. If students applied multiple times and/or to multiple majors, we asked them to rank them according to their preference and took the top-ranked major. As a robustness check, we also used the major associated with the highest income (which differed from the top-ranked major in only 28 % of the cases; see Section 5.3) as dependent variable.

We matched income information from the German Microcensus to the majors. Following Glocker and Storck (2014), we used the 2007-2012 Microcensus waves[7] to obtain estimates of average net hourly earnings reported by graduates in various majors. We first restricted the Microcensus sample to individuals below 65 years of age holding a tertiary degree (universities and universities of applied science). Furthermore, we only included those whose own labor income was their primary source of income. The Microcensus measures net

[6] A Google search of "college major & income [Studienfach & Einkommen]" produces around 140,000 hits. Several websites contain more or less detailed information on the average wages of college graduates by college major (mainly net or gross hourly or annual wages, often for career entrants). A comparably detailed and informative source is the website of the news magazine "Spiegel Online," which provides a search tool on mean net hourly and annual income based on Glocker and Storck (2014).

[7] Sources: DOI: 10.21242/12211.2007.00.00.3.1.0 - DOI: 10.21242/12211.2012.00.00.3.1.0

monthly earnings in 24 categories.[8] We converted this measure into a continuous variable using the categories' midpoints. Using the German Socio-Economic Panel, Glocker and Storck (2014) show that these midpoints are good estimates of mean wages within the categories. Finally, we calculated hourly wages by dividing monthly income by monthly work hours and adjusted the wages for inflation using the German consumer price index. From this data set, we calculated average net hourly wages by major and discarded majors with fewer than 100 observations (for a similar approach, see Ochsenfeld, 2016). This yielded data for 70 majors (see Appendix Table A2). In our treatment, we presented only a subsample of these majors (8 out of 70). Showing all majors would not have been feasible (particularly when comparing them to apprenticeship occupations). As explained earlier, the aim of our treatment was not to provide exhaustive information but rather to make students aware of possible misconceptions and information deficits.

In our analytical sample, average hourly earnings range from 12.13 € to 25.77 € (mean: 17.2; SD: 2.58). Note that we did not differentiate our income measure by gender. Despite substantial earning differentials, the relative ranking of majors is fairly similar for men and women. Furthermore, we are interested in capturing "the true enduring value of different fields" (Davies & Guppy, 1997, p. 1424). Importantly, this calculation is also closer to our treatment, as we did not provide information on gender-specific returns to majors. As our dependent variable does not account for part-time employment, which is overrepresented in female-dominated occupations, we conducted a further robustness check with (a) net monthly income of all (self-)employed and (b) net monthly income of (self-)employed working full-time (more than 35 h a week) as the dependent variables. The results remain unchanged.

Our dependent variable differs strongly from most other studies on gendered major choices, which categorize majors into STEM vs. others (e.g., Lörz et al., 2011; Mann & DiPrete, 2013; Morgan et al., 2013; Zafar, 2013) or into a wider range of categories (e.g., Barone, 2011; Jonsson, 1999). Our approach has at least two advantages: First, variance

[8] Unfortunately, the Microcensus does not provide information on gross income. This may lead to an underestimation of income in female-dominated fields due to German tax regulations. Most married couples file their taxes jointly, leading to lower tax rates on total household income but consequently higher tax rates for minor incomes within the household, which more often affects women.

within one major category is large with regard to earning prospects and other dimensions.[9] Thus, broad categorical measures of majors run the risk of missing substantial income differences that might occur when comparing students within a broader major category. Second, and importantly, our measure captures the hierarchical nature of major choices regarding one important dimension—earnings—that is closely linked to our treatment, in which we provided information on monetary returns to different majors (see Section 4.4).[10] To compare our results with those of Barone and colleagues (2019), we also rerun our analyses using their categorization. The results remain substantially the same but miss statistical significance (see Section 5.3).

Our two main *independent* variables are treatment status and gender. Moreover, we include job attribute preferences to test our hypotheses 2- 4. In the first wave, students were asked to indicate the relevance of several attributes to their occupational choices on a four-point scale (ranging from very important to not at all important). We selected four attributes that represent gender-typical roles and stereotypes introduced in Section 2. Two refer to gender roles—“high income” and “sufficient time to meet family obligations”—and two to gender stereotypes—“opportunity to help other people” and “having much contact to other people.” We chose to construct four dummy variables differentiating between “very important” and the remaining categories because this is the category that better differentiates between male and female students. Furthermore, due to the sample size, including more categories was hardly possible for some of our tests (especially H3a/b and H4a/b).

4.6 Analytical strategy

With perfect randomization, we would only need *t-tests* for outcome differences between the TG and the CG to identify average treatment effects (ATE). However, as is usually done in

[9] For instance, studying medicine leads to a net hourly income of 23.23 € on average, whereas graduates of “health studies” only earn 15.65 € on average. A second example is physics (19.33 €) and biology (15.88 €).

[10] Note that earnings also vary within majors. Some majors have high average earnings but also high variance, indicating that some graduates will earn much less and some much more. Interestingly, level and variance of field-specific earnings do not correlate, indicating that both high- and low-paying majors have high variances (Glocker & Storck, 2014). However, we do not consider this issue further in this paper, because our treatment did not mention variance; neither do publicly available information materials include information about this.

the field of education, we randomized at school level to avoid contamination bias and to simulate a feasible policy measure. Cluster randomization, however, makes imperfect covariate balance between treated and control students at the individual level more likely because students' selection into schools is not completely random and the number of schools is rather small in our study, making random draws with unbalanced covariates likely. Furthermore, students' participation was of course voluntary, which might lead to selective participation and panel attrition based on school-specific characteristics.

First, we did not observe selective attrition between experimental groups (see Table A1 in the Appendix). Second, we checked the co-variate balance between TG and CG for the whole analytical sample and within the two gender categories for a large number of variables that might relate to (gender differences in) major choices and that might influence the way students process and respond to the information treatment (e.g., their pre-treatment feeling of being informed about HE and VET or the information sources they used before). Table A3 in the Appendix identifies some differences between students in the CG and TG. Some of the percentage point differences appear substantial, which is also due to small case numbers, especially within the male and female samples. We therefore adjust distributional differences by using a reweighting strategy that includes all variables that show a significant difference at the 10 percent level either in the overall sample or the gender-specific samples (job attribute preference: high income, locus of control, number of information sources used privately, feeling informed about VET) or that are central to our hypotheses (further job attribute preferences). Our reweighting does not address possible unobserved differences between the groups. Yet in contrast to survey data, we randomly assigned the treatment, thereby already excluding some unobserved differences by design. In addition, we searched quite extensively for theoretically meaningful observable differences (which often approximate unobservable characteristics).

The basic idea of the reweighting is to address the question: What would be the treatment effect if there were no distributional differences between TG and CG? We therefore weight group CG in such a way that it matches the distribution of several variables x of group TG. To account for gender-specific imbalances, we also added interactions with gender for all variables. For obtaining the weights that "balance" the two groups with respect to the set of variables x, we use *entropy balancing* (Hainmueller, 2012). This technique reweights our data so that means and higher moments of a variable are matched in the two groups. It is especially appropriate for our small sample because the issue of not covering "statistical twins" does not arise, and it is more effective than parametric approaches such as propensity score matching.

We also use entropy balancing to test whether gendered job attribute preferences mediate the potential gender difference in the treatment effect (H2). For this, we stepwise reweight the group of male students so that they equal female students with regard to their job attribute preferences. Afterwards, we compare the treatment effect for the adjusted male students with their "original" treatment effect. A significantly reduced treatment effect would support our mediation hypothesis H2. To test our moderation hypotheses 3a/b and 4a/b, we finally estimate three-way interaction effects between the job attribute preferences, gender, and the treatment. To do so, we additionally complement the balancing with the interaction terms between the job attribute preferences and the balancing variables. We then estimate the treatment effects for men and women separated by their stated preferences. We prefer this method over regression-based decomposition techniques for two reasons: First, it is a natural extension of the method used for addressing imbalances between TG and CG, as described above, and it enhances the consistency of our analyses. Second, it is non-parametric, meaning we do not need to assume functional forms for the relationships between the variables.

To test for differences between TG and CG, we use *two-tailed t-tests* for clustered data. Standard methods to calculate the *t* statistic assume independently sampled individuals. This is not the case in our analysis, because the students are nested in 27 schools. However, standard methods to deal with this type of clustering are biased if the number of clusters is smaller than 50 (Cameron & Miller, 2015). To avoid bias, we estimate the p-values for the t-tests using wild cluster bootstrap with equal weights, as proposed by Cameron, Gelbach, and Miller (2008).[11]

5 Findings

5.1 Treatment effect on male and female students' major choices

As a starting point, we look at average incomes in the fields students intend to study one year prior to graduation and thus before the treatment. Figure 2 shows the distribution of major-specific incomes by gender. The kernel density plots clearly show that male students plan to

[11] We implemented this using the user-written Stata ado "cgmwildboot" by Judson Caskey, available here: https://sites.google.com/site/judsoncaskey/data.

study higher-paying fields than female students. This is also reflected in a mean difference of about 1€ per hour (p<0.001).

[**Figure 2**]

The first row of Table 1 shows the average treatment effect calculated as the difference in major-specific hourly net income between TG and CG at the time of application. The results indicate that the treatment did not have an overall effect (first columns) but influenced the major choice of male (but not female) students. Treated men apply to majors that pay on average 1€ more than majors chosen by otherwise similar men in the CG. This difference is statistically significant. In contrast, for treated women the effect is not significant, if notable at all; they even apply to slightly less well-paying majors than women in the CG. The last column (first row) shows that the gender difference in the treatment effect is substantial and significant. These findings are in line with the prediction of H1.

[**Table 1**]

The finding of a larger treatment effect among men is also backed by the results in the second row of Table 1, which displays individual-level *changes* between the intended major (pre-treatment) and the major applied to (post-treatment). This strategy additionally controls for heterogeneity between the experimental groups due to potential unobserved and time-constant individual- and school-level confounders. At the same time, this analysis only considers *changes* in major-specific wages. Thus, it neglects the potentially stabilizing effect of the treatment on remaining in a well-paid major. Nevertheless, the results remain qualitatively similar—a larger treatment effect among men—, which corroborates our findings in the first row of Table 1. Corresponding to the fact that this method captures only a part of the treatment effect, the effect size becomes smaller among men.[12] Also, the effect among men is not statistically significant. This is not surprising, given the reduced variance in the dependent variable (one third of the sample remains stable (see Table 2)). Because of these limitations,

[12] Indeed, our data show that treated male students with stable intentions apply to majors associated with a higher income than otherwise similar students in the CG.

we only use the changes within individuals presented in the second row of Table 1 as a sensitivity analysis and return to the comparisons of levels as displayed in the first row of Table 1 to test the remaining hypotheses.

To gain more insights into the pattern underlying the average treatment effects, Table 2 displays descriptive information on the proportion of stability and changes between intended major and major applied to. Looking at any change, about one third of our male and female respondents in both TG and CG did not change their major (in terms of income). Moreover, this analysis reveals that the positive treatment effect for men compared to women is mainly generated by fewer downward moves in the TG (only 27 % compared to 42 %), whereas treated women experienced more downward moves (though not statistically significant). The differences are less pronounced for more substantial changes (at least +/- 1€).[13]

[**Table 2**]

5.2 Mediation and moderation

Next, we turn to the mediating influence of job attribute preferences (H2). Table 3 generally supports the idea of gender differences in these preferences: Whereas male students more often report that high income is important to them, female students place much greater value on social contacts and helping people. Interestingly, the difference in the preference for a job that allows for reconciling family and work ("time for family") is small.

[**Table 3**]

Table 4 shows that gender differences in our measured job attribute preferences explain only a small part of the gender difference in treatment effect. The baseline difference in the first row is taken from the first row in Table 1. Starting with the second row, we reweighted the sample

[13] Examples of very substantial changes involve upgrades to or downgrades from medicine and dentistry (most often from/to biology); small changes include switches both within (e.g., between German studies and history) and across (e.g., between economics and psychology) a broader major category.

of men so that it resembles the distribution of women's preferences. Adjusting for compositional gender differences in the preference for high income (2nd row of Table 4) does not reduce the gender difference in treatment effects. The same is true if we additionally adjust for preferences for "time for family" or "helping others" (3rd and 4th row). Only the combination of the motives "high income" and "contact with others" (5th row) reduces the gender difference in the treatment effect. This reduction, however, only amounts to about 0.14 €, which is negligible compared to the baseline treatment effects. Furthermore, the bootstrapped p-value is far from statistically significant. Thus, distributional differences in job attribute preferences do not seem to be the mechanism behind the gender difference in the treatment effect. H2 is therefore not supported.

[**Table 4**]

Turning now to our moderation hypotheses, the results in Table 5 suggest that preferring jobs that lead to a high income increases the treatment effect among male students (to 1.48 €). Among female students with the same stated income preferences, the ATE is close to zero. As a consequence, the gender difference in the ATEs among those who prefer a high income is 0.21 € higher compared to those who do not state this preference (last column). Although this difference is at odds with hypothesis H3b, it points into the direction suggested by H3a, which states that the income motive is more salient for men, thus making them more responsive to monetary information. Yet the difference is small and not statistically significant. We therefore conclude that the explanatory power of the income preference for the gender gap in the treatment effect is limited at best.

[**Table 5**]

Regarding "female-typical" job attributes, Table 5 further shows that preferring a job that allows "time for family" does not change the ATE among female students (-0.41 € in Table 1 and -0.36 € in Table 5). However, it substantially *decreases* the ATE among male students to only 0.65 € (in Table 5). As a consequence, the gender difference in the ATE (column 3) is substantially reduced to only 1 €. Moreover, this difference is smaller than the gender difference in the ATE among those for whom “time for family” is less important, although the “difference-in-difference” of 0.34 € is not statistically significant (last column).

High rates for “helping others” substantially increase the treatment effect for both male and female students—to 1.98 € and 0.55 € (though not statistically significant for women), respectively—however, the gender difference remains unchanged (1.43 €). Compared to those men and women who do not consider “helping others” important, the difference is again slightly, but not significantly, smaller. The last point is also true of the female-stereotypical preference for jobs that allow for “social contact with others.”

With regard to hypotheses H4a and H4b, the findings are clearly at odds with the former, thus rejecting the assumption of a *stronger* “male advantage” within the group of students who state preferences for “female-typical” job attributes. Yet the findings referring to the gender role-related job attribute “enough time for family” point in the direction of H4b (suggesting the “male advantage” to be weaker within this group of students). Yet the “difference-in-difference” is again not statistically significant, leading us to reject H4b.

5.3 Robustness checks

To test the robustness of our results, we conducted a series of additional analyses. The results are reported in Table 6. Including respondents with missing information on their intended major, and those who did not intend to enroll in college in wave 1, supports our findings of a positive treatment effect for male students. However, if we use the major associated with the highest income (among those majors each student applied to) instead of the first-ranked major as dependent variable, the ATE for male students is reduced to 0.44 €. The explanation for this diverging finding is that the “highest-income” major and the “top-ranked” majors are more often the same among treated than untreated male students (82 % vs. 73 %, not shown in the table). Thus, for male students the intervention apparently increased both the average income-return associated with major choices *and* the correspondence between the major leading to the highest income and the preferred one. For female students, we again do not observe such differences, as the match between preferred and highest-income major is the same in the TG and the CG (68 % and thus interestingly lower than for their male peers). Using monthly instead of hourly wages (and thereby accounting for differences in average working hours between occupations) corroborates our finding.

Furthermore, replicating the broader categories used by Barone and colleagues (2019), we observe that the treatment increases the share of male students applying to strong fields by 9 percentage points and decreases the share of male students applying to weak fields by 6 percentage points. The share of female students applying to fields in the three categories is, however, unaffected by the treatment. Thus, even with a more similar operationalization of major choices, our findings are in conflict with those reported by Barone and colleagues (2019). This corroborates our explanation that the different treatments (and not the different operationalization of major choice) might drive the differences in the results. The described treatment effect for male students is not significant when using these broader major categories. One reason for this is that quite substantial differences in major-specific wages within one broader category—for example, between construction engineering (15.60 €), mechanical engineering (18.44 €), and dentistry (25.77 €), all belonging to the "strong field"—are disregarded and thus do not contribute to the treatment effect (see Table A2 in the Appendix).

[**Table 6**]

Finally, one might argue that actual *enrollments* are more relevant for labor market opportunities than applications. We therefore rerun our analysis with the average income associated with the major enrolled in as the dependent variable. The finding is the same: Men benefitted more from the information treatment than women, although the gender differences are smaller, possibly due to the impact of admission decisions by the HE institutions.[14]

To summarize, we find a quite robust gender gap in the treatment effect, but differences in the salience of gendered job attribute preferences only explain a small part of it: A substantial gender gap remains even among men and women with similar preferences. Yet we do find some, albeit uncertain, indications that preferences for certain job attributes seem to be more salient for male than for female students in that they alter especially men's responsiveness to information on income returns. Interestingly, this is the case for both job attributes that are in line with the common male breadwinner role ("high income") *and* for those that deviate from it ("time for family"). There are no indications, however, that job

[14] The gender difference in the ATE is 0.91 € ($p < 0.1$); the difference between treated und non-treated women is -0.13 € ($p > 0.1$) and 0.72 € ($p < 0.05$) between treated und non-treated men.

attribute preferences that are linked to female-typical stereotypes ("helping others," "contact with others") alter the treatment effect in a gender-specific way. This suggests that further explanations for the gender gap in the income potential of the major(s) applied to might, for instance, be found in the social costs of choosing gender-atypical majors or in reflections on gender-specific comparative advantages in different fields (e.g., because of differences in course work) (Jonsson, 1999), which are likely linked to different perceptions of requirements for admission and graduation.

6 Discussion and conclusion

In this paper, we investigated whether information on major-specific income returns to HE influences students' major choices in a gender-specific way. We hypothesized that this might indeed be the case because of gender stereotypes and gender roles, which influence related preferences for certain attributes of future jobs and ultimately gender differences in the "in/attentiveness" to monetary information on college majors. We studied this research question by employing an experimental design with an information treatment on returns to HE in Germany.

Our main findings are: First, additional information on income differences between college majors (treatment) is only influential for male but not for female students. On average, only treated men applied more often to majors leading to a higher income than untreated men. Second, although the distribution of job attribute preferences differs strongly by gender, the gender difference in the treatment effect cannot be explained by this distributional difference (mediation). Third, concerning gender-stereotypical job attribute preferences, we do not find a differential treatment effect with regard to the importance of having "much contact with others"—neither for male nor for female students. However, for those who rated "helping others" as very important, we find substantial increases in the treatment effect for both men and women. The latter could be due to the comparably high-income returns of some majors in the care-related fields (e.g., psychology, medicine, or teaching). Medicine in particular seems to combine different characteristics considered to be more or less attractive for male and female students alike (Morgan et al., 2013): Medical jobs involve a high income and occupational prestige; they often require unfavorable working hours, and they are equally oriented towards caring, people, and hard sciences. Labor-market oriented students with preferences for care-related jobs might have listened to our information workshop and realized that it is possible to achieve a high income without necessarily having to abandon their care-related preferences. Fourth, concerning *gender role*-related job attribute preferences, our findings may suggest that such preferences seem to be more salient for male than for female students in that they especially alter men's responsiveness to information on income returns. This is the case for both gender-role-conforming ("high income") *and* non-conforming ("time for family") job attributes. The treatment effect substantially increases among men who rate the "male-typical" job attribute "high income" as very important, whereas it substantially decreases if we consider only men with "female-typical" preferences ("time for family"). Hence, these findings provide first indications that the fit between stated preferences, the use of (in)compatible information, and actual decisions seems to be tighter for male than for female students.

Although the main finding of a substantially stronger treatment effect among male than among female students is in line with cultural accounts, the responsiveness of male students to monetary information (especially if income is very important for their occupational choices) indicates that they actively engage with this information in a deliberative way and integrate it into their decision-making, which is (also) in line with rational choice models. The stronger responsiveness of male students does, however, not necessarily mean that female students decide in a non-rational way. First, as argued in Section 2.3, inattention may also be a rational response if the information is considered irrelevant (McGuigan et al., 2016). Second, male and female students might evaluate the benefits of educational decisions against different criteria, which again are a result of socialization processes and perceived social costs (Jonsson, 1999; Lörz et al., 2011; Ochsenfeld, 2016).

Our German results are different from the findings by Barone and colleagues (2019) for Italy (see Section 2.1). One obvious reason might be that our continuous measure for major choice strongly differs from their categorical measure. However, as our robustness check reveals, our findings remain substantially the same when using their categories (see Section 5.3). We therefore believe that the reason rather lies in the different treatments, which is also supported by the stronger similarity of our findings with those of the Finnish study by Kerr and colleagues (2014). Similar to the Finnish intervention, our information workshop more strongly focused on financial returns, which might have rendered the information less salient for female students. In contrast, the Italian study also included other more gender-neutral employment dimensions like first job search durations. Interestingly, the Finnish study only found a positive treatment effect for boys from less-educated neighborhoods (see Section 2.1). Our study was also conducted in such neighborhoods (see Section 4.1) and thus seems to corroborate this finding. The overall small sample size, however, does not allow us to further examine differential treatment effects by gender and (individual) social background. As the interaction between social background and gender in educational choices is generally understudied, this might be a promising avenue for future research.

What are possible social implications of our findings? This article was motivated by literature on the persistent gender-wage gap and the influential role that gender inequalities in major choices play for that gap. At first glance, the results of our field-experimental data support the pessimistic view that "correcting" students' biased information on monetary returns to fields of study does not help reduce gender differences in income-related major choice—though correct monetary information seems to have the potential to reduce inequality between social classes regarding the decision for or against HE (see Ehlert, Finger et al., 2017; Loyalka et al., 2013; Oreopoulos & Dunn, 2013; Peter & Zambre 2017). Yet we have to keep in mind that our intervention took place late in the school career and was of short duration. Thus, earlier and more intensive interventions might be more successful in this respect (e.g., Barone et al., 2019). Moreover, another result from our study is also more promising: As reported above (Section 4.3), overall our information intervention increased the college application rates of those who intended to apply to college at the end of the penultimate school year but not of those who intended to start an apprenticeship. A closer look at the gender difference reveals that this "stabilization" effect is somewhat more pronounced for women than for men (analysis is not shown in the paper). This means that our intervention did increase the likelihood of women to apply to college instead of doing an apprenticeship (which usually leads to lower wages)—though it did not channel them to highly paid college majors. As Barone and colleagues (2019) suggest, providing information on other employment dimensions than monetary returns might also be promising for equalizing male and female choices of college majors.

Finally, our study has some limitations. First, the data only allow for treating job attribute preferences as single ratings. There are many overlaps between the four dummy variables, and especially the women in our sample rate a higher number of items as "very important," which possibly mirrors their preferences for an adaptive lifestyle (Hakim, 2002). Male students, however, seem to be somewhat firmer in their stated preferences. A second limitation is that the validity of our study is restricted to the context in which it took place, that is, in socially more disadvantaged districts in a major German city. It is possible, for instance, that the treatment effect we detected is somewhat stronger than it would be in socially advantaged districts, as students there can be expected to be somewhat better informed and to have better private (and school-related) information sources, meaning that external information may be less fruitful (Betts, 1996; Hastings et al., 2016). Furthermore, the effect of providing information on returns might be reduced in rural areas, where the supply of HE institutions and majors is clearly limited, meaning that the social and economic costs related to geographic mobility may restrict students' choice set. However, reducing the potential influence of such confounders at different levels helps to identify the underlying mechanisms and effectiveness of a treatment for the group under consideration. Finally, we have to admit the rather small sample size of our study, especially at the school level (the level of our randomization) but also at the individual level. Especially the group of treated men is small; even more so after differentiating by job attribute preferences. This obviously leads to imprecise estimates. Despite this, the estimates among men still reach conventional levels of statistical significance. Apparently, the treatment effect is strong enough to be visible even under such adverse circumstances. With our robustness checks, we showed that our results are stable even with different specifications of the sample and the dependent variable. We are thus confident that the limitations of our data did not lead to biased results. Further experimental studies would, however, benefit from larger sample sizes to increase the robustness of findings but also to allow for investigating effect heterogeneity (e.g., concerning the interaction between gender and social or migration background).

Declaration of Competing Interest

None.

Acknowledgements

This research was supported by a grant from the Einstein Foundation (grant number A-2010-25 “Berliner-Studienberechtigten-Panel / Best Up”). We thank our Best Up colleagues at the DIW Berlin for their productive cooperation in the joined panel data collection (Matthias Hübener, Frauke Peter, C. Katharina Spieß, Johanna Storck, and Vaishali Zambre) and our student assistants (Christina Altmann, Cindy Fitzner, Judith Heinmüller, Torsten Hovorka, Johanne Klindworth, and Fabio Krauthäuser) for their valuable support.

References

Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education*, 84(2),157-176. https://doi.org/10.1177/0038040711402099

Barone, C., Schizzerotto, A., Abbiati, G., & Argentin, G. (2017). Information barriers, social inequality, and plans for higher education. *European Sociological Review*, 33(1), 84-96. https://doi.org/10.1093/esr/jcw050

Barone, C., Schizzerotto, A., Assirelli, G., & Abbiati, G. (2019). Nudging gender desegregation: A field experiment on the causal effect of information barriers on gender inequalities in higher education. *European Societies*, 21(3), 356-377. https://doi.org/10.1080/14616696.2018.1442929

Berger, M. (1988). Predicted future earnings and choice of college major. *Industrial and Labor Relations Review*, 41(3), 418–29. https://doi.org/10.2307/2523907

Betts, J. (1996). What do students know about wages? *Journal of Human Resources*, 31(1), 27–56. https://doi.org/10.2307/146042

Bobbitt-Zeher, D. (2007). The gender income gap and the role of education. *Sociology of Education*, 80(1),1-22. https://doi.org/10.1177/003804070708000101

Böhm, C., & Pampel, J. (2014). *Landesprogramm der vertieften Berufsorientierung für Berliner Schülerinnen und Schüler (BVBO) – Sachbericht 2013/2014*. Berlin: SPI Consult GmbH, https://web.archive.org/web/20160507175204/http://www.spiconsult.de/fileadmin/Dokumente/bildung/BVBO6/Sachbericht13_14.pdf (accessed 2019/07/17).

Boliver, V. (2013). How fair is access to more prestigious UK universities? *The British Journal of Sociology*, 64(2), 344-364. https://doi.org/10.1111/1468-4446.12021

Bradley, K. (2000). The Incorporation of women into higher education: paradoxical outcomes? *Sociology of Education*, 73(1), 1-18. https://doi.org/10.2307/2673196

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317-372. https://doi.org/10.3368/jhr.50.2.317

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Ceci, S.J., Williams, W.M., & Barnett, S.M. (2009). Women's underrepresentation in science. *Psychological Bulletin*, 135(2), 218-261. http://dx.doi.org/10.1037/a0014412

Charles, M., & Bradley, K. (2002). Equal but separate: A cross-national study of sex segregation in higher education. *The American Sociological Review*, 67(4), 573–599. https://doi.org/10.2307/3088946

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? *American Journal of Sociology*, 114(4),924-976. https://doi.org/10.1086/595942

Davies, P., Davies, N. M., & Qiu, T. (2017). Information and choice of A-level subjects. *British Educational Research Journal*, 43(4), 647-670. https://doi.org/10.1002/berj.3289

Davies, S., & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social Forces*, 75(4),1417–1438. https://doi.org/10.2307/2580677

Daymont, T.N., & Andrisani, P.J. (1984). Job preferences, college major, and the gender gap in earnings. *The Journal of Human Resources*, 19, 408–428. https://doi.org/10.2307/145880

DellaVigna, S. (2009). Psychology and economics. *Journal of Economic Literature*, 47(2), 315-372. http://doi.org/10.1257/jel.47.2.315

Denzler, S., & Wolter, S.C. (2010). Der Einfluss des lokalen Hochschulangebots auf die Studienwahl. *Zeitschrift für Erziehungswissenschaften*, 13(4), 683-706. https://doi.org/10.1007/s11618-010-0143-6

Domina, T. (2009). What works in college outreach: Assessing targeted and schoolwide interventions for disadvantaged students. *Educational Evaluation and Policy Analysis*, 31(2), 127–152. https://doi.org/10.3102/0162373709333887

Ehlert, M., Finger, C., Rusconi, A., & Solga, H. (2017). Applying to college. Do information deficits lower the likelihood of college-eligible students from less-privileged families to pursue their college intentions? *Social Science Research*, 67, 193-212. https://dx.doi.org/10.1016/j.ssresearch.2017.04.005

Ehlert, M., Peter, F., Finger, C., Rusconi, A., Solga, H., Spieß, C. K. and Zambre, V. (2017). The Berliner-Studienberechtigten-Panel (Best Up) – Methodological and Data Report. *Data Documentation*, 90. Berlin: DIW. https://www.diw.de/documents/publikationen/73/diw_01.c.561179.de/diw_datadoc_2017-090.pdf)

Glocker, D., & Storck, J. (2014). Risks and returns to educational fields. *Economics of Education Review*, 42, 109–129. https://doi.org/10.1016/j.econedurev.2014.06.004

Hainmueller, J. (2012). Entropy balancing for causal effects. *Political Analysis*, 20(1), 25–46. https://doi.org/10.1093/pan/mpr025

Hakim, C. (2002). Lifestyle preferences as determinants of women's differentiated labor market careers. *Work and Occupations*, 29(4), 428-459. https://doi.org/10.1177/073088802237558

Hanson, S. L. (1994). Lost talent: Unrealized educational aspirations and expectations among US youths. *Sociology of Education*, 67(3), 159–183. https://dx.doi.org/10.2307/2112789

Hastings, J. S., Neilson, C. A., Ramirez, A., & Zimmerman, S. D. (2016). (Un)informed college and major choice. *Economics of Education Review*, 51, 136-151. https://doi.org/10.3386/w21330

Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2015). The effects of earnings disclosure on college enrollment decisions. *NBER Working Paper*, 21300. https://doi.org/10.3386/w21300

Helbig, M., & Nikolai, R. (2015). *Die Unvergleichbaren. Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949*. Bad Heilbrunn: Julius Klinkhardt.

Herbaut, E., & Geven, K. M. (2019). What works to reduce inequalities in higher education? A systematic review of the (quasi-)experimental literature on outreach and financial aid, *World bank policy research working paper*, No. 8802. Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/650601554221255443/pdf/What-Works-to-Reduce-Inequalities-in-Higher-Education-A-Systematic-Review-of-the-Quasi-Experimental-Literature-on-Outreach-and-Financial-Aid.pdf (accessed: 2019/10/17).

Huntington-Klein, N. (2016). (Un)informed college and major choice. *Economics of Education Review*, 53, 159-163. https://doi.org/10.1016/j.econedurev.2016.03.008

Jansen, M., & Liefbroer, A. C. (2006). Couples' attitudes, childbirth, and the division of labor. *Journal of Family Issues*, 27(11), 1487–1511. https://doi.org/10.1177/0192513X06291038

Jonsson, J. O. (1999). Explaining sex differences in educational choice. *European Sociological Review*, 15, 391-404. https://doi.org/10.1093/oxfordjournals.esr.a018272

Kerr, S. P., Pekkarinen, T., Sarvimäki, M., & Uusitalo, R. (2014). *Educational choice and information on labor market prospects: A randomized field experiment.* Manuscript. Online: http://www.demm.unimi.it/extfiles/unimidire/100601/attachment/pekkarinen.pdf (accessed: 2019/06/23).

Kerr, S. P., Pekkarinen, T., Sarvimäki, M., & Uusitalo, R. (2015). Post-secondary education and information on labor market prospects: A randomized field experiment. *IZA discussion paper*, No. 9372. Bonn: IZA. http://ftp.iza.org/dp9372.pdf. (accessed: 2019/10/17).

Kolsto, S. D. (2001). 'To trust or not to trust, …'- pupils' ways of judging information encountered in a socio-scientific issue. *International Journal of Science Education*, 23(9), 877–901. https://doi.org/10.1080/09500690117217

Konrad, A. M., Ritchie Jr, J. E., Lieb, P., & Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: A meta-analysis. *Psychological Bulletin*, 126(4), 593–641. https://doi.org/10.1037/0033-2909.126.4.593

Legewie, J., & DiPrete, T. A. (2014). The high school environment and the gender gap in science and engineering. *Sociology of Education*, 87, 259-280. https://doi.org/10.1177/0038040714547770

Leuze, K., & Strauß, S. (2009). Lohnungleichheiten zwischen Akademikerinnen und Akademikern. *Zeitschrift für Soziologie*, 38(4), 262-281. https://doi.org/10.1515/zfsoz-2009-0401

Lörz, M., Schindler, S., & Walter, J. G. (2011). Gender inequalities in higher education. *Irish Educational Studies*, 30, 179-198. https://doi.org/10.1080/03323315.2011.569139

Loyalka, P., Song, Y., Wei, J., Zhong, W., & Rozelle, S. (2013). Information, college decisions and financial aid: Evidence from a cluster- randomized controlled trial in China. *Economics of Education Review*, 36, 26–40. https://doi.org/10.1016/j.econedurev.2013.05.001

Mann, A., & DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Social Science Research*, 42, 1519-1541. https://doi.org/10.1016/j.ssresearch.2013.07.002

Mayer, K.U., Müller, W., & Pollak, R. (2007). Germany: Institutional change and inequalities of access in higher education. In: Y. Shavit, R. Arum, & A. Gamoran (Eds.), *Stratification in higher education. A comparative study*. Stanford, CA: Stanford University Press, 240–265.

McGuigan, M., McNally, S. & Wyness, G. (2016). Student awareness of costs and benefits of educational decisions: Effects of an information campaign. *Journal of Human Capital*, 10(4), 482-519. http://doi.org/10.1086/689551

Montmarquette, Cannings, K., & Mahseredjian, S. (2002), How do young people choose college majors? *Economics of Education Review*, 21, 543–556. https://doi.org/10.1016/S0272-7757(01)00054-1

Morgan, J. H. M. (2010). *The role of financial information in college decision making*. PhD thesis. Boston, MA: Boston College. Retrieved December 13, 2018 (http://hdl.handle.net/2345/1407).

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, 42(4), 989-1005. https://doi.org/10.1016/j.ssresearch.2013.03.008

National Education Report 2018 = Autorengruppe Bildungsberichterstattung (2018). *Nationaler Bildungsbericht 2018*. Bielefeld: Bertelsmann.

Ochsenfeld, F. (2016). Preferences, constraints, and the process of sex segregation in college majors. *Social Science Research*, 56, 117–132. https://doi.org/10.1016/j.ssresearch.2015.12.008

OECD (2018). *Education at a glance 2018: OECD indicators*. Paris: OECD Publishing.

Oreopoulos, P., & Dunn, R. (2013). Information and college access: Evidence from a randomized field experiment. *The Scandinavian Journal of Economics*, 115(1), 3–26. https://doi.org/10.1111/j.1467-9442.2012.01742.x

Peter, F., Rusconi, A., Solga, H., Spieß, C. K., & Zambre, V. (2016). Informationen zum Studium verringern soziale Unterschiede bei der Studienabsicht von AbiturientInnen. *DIW-Wochenbericht*, 26, 555-565.

Peter, F., & Zambre, V. (2017). Intended college enrollment and educational inequality. *Economics of Education Review*, 60, 125-141. https://doi.org/10.1016/j.econedurev.2017.08.002

Polachek, S.W. (1981). Occupational self-selection. A human capital approach to sex differences in occupational structure. *The Review of Economics and Statistics*, 63(1), 60-69. https://doi.org/10.2307/1924218

Powell, J.J.W., & Solga, H. (2011). Why are higher education participation rates in Germany so low? *Journal of Education and Work*, 24(1-2), 49-68. https://doi.org/10.1080/13639080.2010.534445

Saniter, N., Schnitzlein, D.D., & Siedler, T. (2019). Occupational knowledge and educational mobility: Evidence from the introduction of job information centers. *Economics of Education Review*, 69, 108-124. https://doi.org/10.1016/j.econedurev.2018.12.009

Spangenberg, H., Mühleck, K., & Schramm, M. (2012). Erträge akademischer und nicht-akademischer Bildung. *HIS: Forum Hochschule*, 11/2012. Hannover: HIS.

Spieß, C.K., & Wrohlich, K. (2010). Does distance determine who attends a university in Germany? *Economics of Education Review*, 29(3), 470-479. https://doi.org/10.1016/j.econedurev.2009.10.009

Stocké, V., Blossfeld, H.-P., Hönig, K., & Sixt, M. (2011). Social inequality and educational decisions in the life course. In H.-P. Blossfeld, H.-G. Roßbach, & J. von Maurice (Eds.), *Education as a lifelong process. The German national educational panel study. Zeitschrift für Erziehungswissenschaft, Sonderheft 14/2011* (pp. 103-119). Wiesbaden: VS Verlag für Sozialwissenschaften.

Von Hippel, P.T. (2007). Regression with missing ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37(1), 83-117. https://doi.org/10.1111/j.1467-9531.2007.00180.x

Wiswall, M., & Zafar, B. (2015a). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2), 791–824. https://doi.org/10.1093/restud/rdu044

Wiswall, M., & Zafar, B. (2015b). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2), 117–169. https://doi.org/10.1086/681542

Young, R., & Johnson, D. R. (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and the Family*, 77(1), 277-294. https://doi.org/10.1111/jomf.12144

Zafar, B. (2013). College major choice and the gender gap. *The Journal of Human Resources*, 48, 545-595. https://doi.org/10.3368/jhr.51.2.0713-5837R

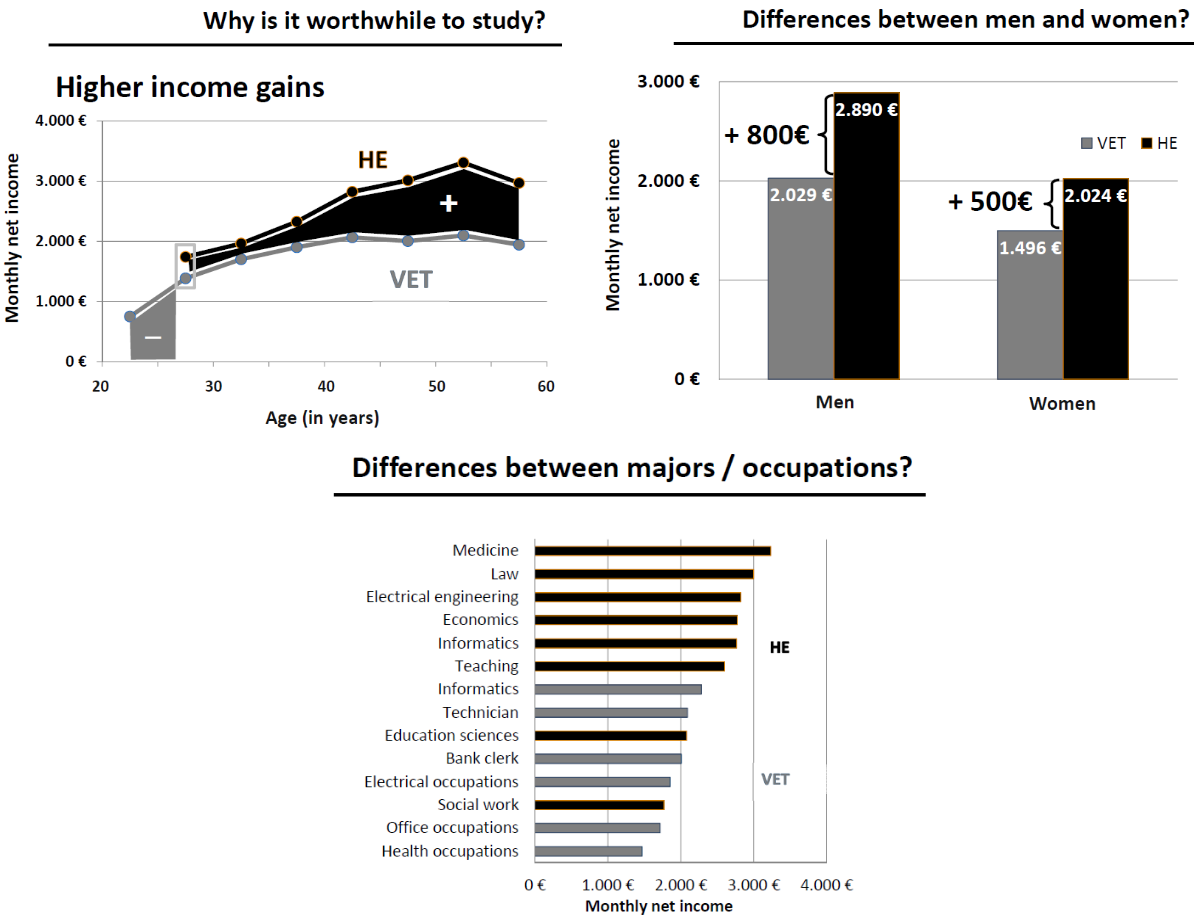
Appendix

[Table A1]

[Table A2]

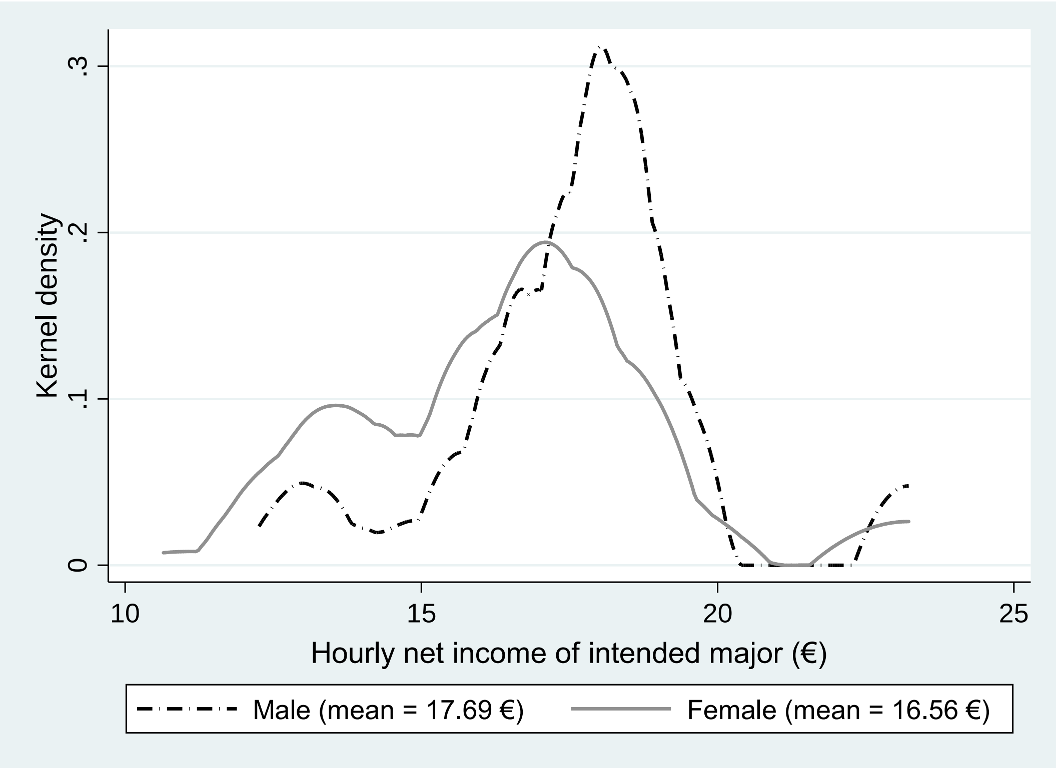
[Table A3]

Figure 1: Sample slides of the treatment with information on income returns



Source: Best Up presentation. Authors' translation.

Figure 2: Distribution of average hourly net income of respondents' intended major, by gender



Sources: Best Up, wave 1 (pre-treatment), analytical sample (see Section 4.3); German Microcensus 2007-2012**.**

Table 1: ATE on average hourly net income of major applied to (in €)

Dependent variable	All			Women			Men			Gender diff. in ATE (p-value)
	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	
Average income of major applied to	17.26	17.08	0.18 (0.668)	16.55	16.96	-0.41 (0.368)	18.29	17.25	1.03* (0.012)	1.44** (0.008)
Individual-level change: av. income of major intention and application	0.18	0.19	-0.01 (0.954)	-0.05	0.55	-0.59 (0. 108)	0.51	-0.33	0.84 (0.160)	1.44 (0. 114)
N		510			307			203		510

Notes: Adjusted for differences between CG and TG (overall and within gender categories): job attribute preferences, locus of control, number of information sources used privately, feeling informed about VET. p-values based on wild cluster bootstrap: + p < 0.1, * p < 0.05, ** p < 0.01.
Sources: Best Up, waves 1-5, analytical sample (see Section 4.3); German Microcensus 2007-2012.

Table 2: Income dynamics between major intended and major applied to (in %)

Dependent variable	All			Women			Men			Gender diff. in ATE (p-value)
	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	
Any change										
Down	32	33	-1 (0.84)	35	26	8 (0.22)	27	42	**-15*** (0.05)	**-23*** (0.03)
Stable	31	32	-1 (0.82)	31	35	-4 (0.48)	32	28	4 (0.55)	08 (0.41)
Up	37	35	2 (0.60)	35	39	-4 (0.49)	41	30	11 (0.35)	15 (0.32)
At least 1 €										
Down	23	24	-1 (0.84)	26	22	4 (0.54)	20	28	-8 (0.23)	-12 (0.27)
Stable	52	50	2 (0.75)	47	48	0 (0.95)	58	53	4 (0.63)	4 (0.68)
Up	25	25	-1 (0.87)	26	30	-4 (0.29)	23	19	4 (0.66)	8 (0.51)
N		510			307			203		510

Notes: Adjusted for differences between CG and TG (overall and within gender categories): job attribute preferences, locus of control, number of information sources used privately, feeling informed about VET.
Deviation from ATE = Mean(TG) – Mean(CG) due to rounding
p-values based on wild cluster bootstrap: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
Sources: Best Up, waves 1-5, analytical sample (see Section 4.3); German Microcensus 2007-2012.

Table 3: Gender difference of job attribute preferences

Very important for occupational choice:	Women (N=307)	Men (N=203)	Difference (p-value)
a) High income	32 % (98)	41 % (84)	-9^{+} (0.052)
b) Time for family	45 % (137)	41 % (84)	4 (0.292)
c) Contact with others	36 % (109)	19 % (39)	16^{**} (0.001)
d) Helping others	33 % (101)	22 % (44)	11^{*} (0.012)
Only high income	9 % (29)	18 % (36)	-8^{**} (0.002)
At least one of a-d	70 % (215)	55 % (111)	15^{**} (0.006)

Note: p-values based on wild cluster bootstrap: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
Source: Best Up, wave 1, analytical sample (see Section 4.3).

Table 4: Differences in male and female ATE, adjusted for gender differences in job attribute preferences

Very important for occupational choice:	Gender diff. in ATE (p-value)	Difference to base model (p-value)
Base model (Table1)	1.44**	
	(0.002)	
Adjusted for gender differences in job attribute preferences		
High income	1.44**	0
	(0.008)	(.)
High income + time for family	1.40**	-0.04
	(0.006)	(0.694)
High income + helping others	1.52**	0.08
	(0.002)	(0.727)
High income + contact with others	1.30*	-0.14
	(0.034)	(0.549)
All	1.35**	-0.09
	(0.012)	(0.789)

Notes: Adjusted for differences between CG and TG (overall and within gender categories): job attribute preferences, locus of control, number of information sources used privately, feeling informed about VET. P-values based on wild cluster bootstrap: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
Sources: Best Up, waves 1-5, analytical sample (see Section 4.3); German Microcensus 2007-2012..

Table 5: ATE by gender and job attribute preferences

Job attribute preferences	Very Important	ATE Women (p-value)	ATE Men (p-value)	Difference of ATEs (p-value)	Diff.-in-Diff. of ATEs (p-value)
High income	Yes	-0.13	1.48^{+}	1.61	
		(0.860)	(0.070)	(0.172)	0.21
	No	-0.55	0.85	1.40^{**}	(0.830)
		(0.290)	(0.114)	(0.006)	
Time for family	Yes	-0.36	0.65	1.01^{+}	
		(0.320)	(0.256)	(0.076)	-0.34
	No	-0.05	1.30^{+}	1.35^{+}	(0.650)
		(0.958)	(0.058)	(0.056)	
Helping others	Yes	0.55	1.98^{+}	1.43	
		(0. 374)	(0.056)	(0.120)	-0.18
	No	-0.77^{+}	0.84	1.61^{**}	(0.864)
		(0.094)	(0.144)	(0.002)	
Contact with others	Yes	-0.30	0.96	1.27	
		(0.702)	(0.470)	(0.344)	-0.17
	No	-0.43	1.01^{*}	1.44^{*}	(0.910)
		(0.342)	(0.016)	(0.040)	
N		307	203	510	

Note: Adjusted for differences between CG and TG (overall and within gender and the respective job attribute preference categories): job attribute preferences, locus of control, number of information sources used privately, feeling informed about VET. p-values based on wild cluster bootstrap: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.
Sources: Best Up, waves 1-5, analytical sample (See section 4.3); German Microcensus 2007-2012.

Table 6: Robustness checks

	All			Women			Men			Gender diff. in ATE (p-value)
	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	Mean TG	Mean CG	ATE (p-value)	
Further specification of the analytical sample:										
				Dependent variable: av. hourly income of major applied to (in €)						
including resp. with missing information on intended major (wave 1)	17.24	16.11	0.13 (0.73)	16.49	16.99	-0.50 (0.216)	18.30	17.28	1.02* (0.01)	1.52** (0.008)
N		540			322			218		540
Including resp. without college intentions (wave1)	17.16	16.95	0.21 (0.508)	16.55	16.86	-0.31 (0.424)	18.05	17.08	0.97** (0.006)	1.28* (0.012)
N		636			378			258		636
Further specification of the dependent variable:										
				Dependent variable: av. hourly income of major applied to (in €)						
"Highest-income" major applied to	17.91	17.93	-0.02 (1.00)	17.47	17.80	-0.33 (0.464)	18.55	18.11	0.44 (0.306)	0.77+ (0.058)
N		510			307			203		510
				Dependent variable: av. monthly income of major applied to (in €)						
Monthly net income, all employees	3005.41	2919.16	86.25 (0.466)	2816.95	2866.75	-49.81 (0.678)	3276.69	2994.44	282.25* (0.014)	332.06* (0.012)
Monthly net income, only full-time employees (> 35 hours/ week)	3256.55	3177.21	79.34 (0.498)	3097.02	3151.82	-54.80 (0.652)	3486.16	3213.50	272.66 * (0.016)	327.46* (0.014)
N		510			307			203		510
Replication of broader categoriesa) (Barone et al., 2019)				*Dependent variable: categorical distribution of majors applied to (in %)*						
Weak field	13	16	-3 (0.55)	19	20	-1 (0.84)	4	10	-6 (0.21)	-5 (0.47)
Intermediate field	55	57	-2 (0.73)	57	58	-1 (0.92)	53	56	-3 (0.64)	-2 (0.87)
Strong field	32	27	5 (0.53)	24	22	2 (0.76)	42	33	9 (0.25)	7 (0.48)
N		510			307			203		510

Notes: Adjusted for differences between CG and TG (overall and within gender categories): job attribute preferences, locus of control, number of information sources used privately, feeling informed about VET.
a) Deviation from 100% due to rounding.
p-values based on wild cluster bootstrap: + p < 0.1, * p < 0.05, ** p < 0.01.
Sources: Best Up, waves 1-5, analytical sample (see Section 4.3), except for "further sample specification";

German Microcensus 2007-2012.

Appendix A

Table A1: Selectivity of analytical sample

Variables	Total (N: 1142)	Included % (N: 510)	Excluded % (N: 632)	Excluded due to: No information on major intention % (N: 82)	No application % (N: 153)	Panel attrition % (N: 256)	Item non-response* % (N: 141)
Treatment	1142	31.57 (161)	29.43 (186)	20.73 (17)	30.07 (46)	32.81 (84)	27.66 (39)
Gender: Female	1135	60.20 (307)	53.12 (332)	55.00 (44)	55.56 (85)	49.21 (124)	56.43 (79)
Job attribute preference: very important							
High income	1131	35.69 (182)	36.39 (226)	33.33 (27)	32.89 (50)	39.92 (101)	35.56 (48)
Time for family	1134	43.33 (221)	45.51 (284)	46.34 (38)	43.42 (66)	42.52 (108)	52.94 (72)
Contact with others	1135	29.02 (148)	32.00 (200)	35.37 (29)	32.03 (49)	30.59 (78)	32.59 (44)
Helping others	1130	28.43 (145)	27.10 (168)	29.27 (24)	25.00 (38)	25.49 (65)	31.30 (41)
Information-related measures							
Classes about “Studying and Occupational career”	1141	60.00 (306)	67.19 (424)	76.83 (63)	64.05 (98)	67.97 (174)	63.57 (89)
Feeling informed about college	1137						
Poorly	330	28.82 (147)	29.19 (183)	33.33 (27)	25.00 (38)	29.80 (76)	30.22 (42)
Partly	378	31.18 (159)	34.93 (219)	32.10 (26)	36.18 (55)	34.90 (89)	35.25 (49)
Well	429	40.00 (204)	35.89 (225)	34.57 (28)	38.82 (59)	35.29 (90)	34.53 (48)
Feeling informed about VET	1095						
Poorly	378	34.51 (176)	34.53 (202)	40.26 (31)	26.00 (39)	39.09 (95)	32.17 (37)
Partly	377	33.92 (173)	34.87 (204)	33.77 (26)	38.67 (58)	34.98 (85)	30.43 (35)
Well	340	31.57 (161)	30.60 (179)	25.97 (20)	35.33 (53)	25.93 (63)	37.39 (43)
Mean (SD) of number of sources used at school	510	4.87 (2.67)	5.42 (2.69)	6.36 (2.42)	6.26 (2.06)	6.48 (2.08)	6.07 (2.04)
Educational achievement / skills measures							
Mean (SD) of grade point average	504	3.03 (0.66)	3.34 (0.65)	3.31 (0.64)	3.37 (0.66)	3.36 (0.67)	3.29 (0.60)
Mean (SD) of verbal competence	510	0.14 (0.98)	-0.13 (1.03)	-0.04 (1.09)	-0.08 (1.05)	-0.14 (0.99)	-0.22 (1.05)
Mean (SD) of figural competence	510	-0.01 (1.02)	-0.15 (1.04)	-0.30 (1.18)	-0.06 (0.97)	-0.27 (1.05)	-0.09 (0.97)
Further variables							
At least 1 parent with academic degree	1124	43.14 (220)	39.25 (241)	41.77 (33)	35.33 (53)	43.95 (109)	33.58 (46)
Migration background	1123	49.51 (251)	57.95 (357)	62.03 (49)	42.11 (64)	63.45 (158)	63.24 (86)
Mean (SD) of external locus of control	510	-0.02 (1.02)	-0.02 (1.00)	0.29 (0.97)	-0.25 (0.94)	-0.01 (1.06)	0.03 (0.91)

School type	1142						
"Gymnasium"	359	34.31 (175)	29.11 (184)	31.71 (26)	26.14 (40)	28.91 (74)	31.21 (44)
Comprehensive school with a "Gymnasium" track	424	35.88 (183)	38.13 (241)	46.34 (38)	36.60 (56)	38.67 (99)	34.04 (48)
Vocational "Gymnasium"	359	29.80 (152)	32.75 (207)	21.95 (18)	37.25 (57)	32.42 (83)	34.75 (49)

Source: Best Up, waves 1-5, general sample restriction: only students with HE intention (wave 1)
*Includes 94 cases with missing information on major applied for (dependent variable) and 47 cases with missing values on further variables included in the analyses.

Table A2: Average hourly net earnings of HE graduates in different majors

	Fields of study	Hourly wages (€)		Fields of study	Hourly wages (€)
1	Dentistry	25.77	36	Classical Languages	15.16
2	Medicine	23.23	37	Music/Musicology	15.13
3	Chemistry	19.45	38	Geosciences/Geography	14.97
4	Physics/Astronomy	19.33	39	Slavic/Baltic/Finno-Ugrian Studies	14.70
5	Law	19.21	40	Marketing	14.70
6	Mathematics/Statistics	18.73	41	Architecture/Urban Planning	14.55
7	Supply Engineering	18.54	42	History	14.53
8	Mechanical Engineering	18.44	43	Nutrition Sciences	14.52
9	Industrial Engineering	18.23	44	Philosophy	14.52
10	Business & Administration	17.96	45	English Studies	14.44
11	Transport Engineering	17.86	46	Agricultural Engineering	14.37
12	Accounting	17.56	47	Educational Science	14.26
13	Computer Science	17.27	48	Linguistics/Economics/Culture	14.24
14	Natural Sciences/Engineering	17.25	49	Social Sciences	14.12
15	Transport	17.24	50	Other Religious Studies	14.10
16	Electrical Engineering	17.23	51	Sport	14.10
17	Economics	17.16	52	Protestant Theology & Religious Studies	13.93
18	Finance and Insurance	17.15	53	Home Economics	13.80
19	Chemical Engineering	17.15	54	Environmental Sciences	13.77
20	Precision Engineering	16.92	55	Non-European Languages and Cultures	13.60
21	Pharmaceutics	16.84	56	Design/Interior Architecture	13.49
22	Psychology	16.80	57	Catholic Theology & Religious Studies	13.48
23	Teaching	16.64	58	Horticultural Sciences	13.45
24	Management Science	16.00	59	Journalism/Media Studies	13.35
25	Biology/Biochemistry/-technology	15.88	60	Communication and Media Engineering	13.22
26	(Public) Security & Order	15.87	61	Philology	13.15
27	Romance Studies	15.76	62	Art Studies	13.05
28	Health sciences	15.65	63	Archival Studies/Library; Management / Documentation Studies	12.95
29	Construction Engineering	15.60	64	Social Work	12.89
30	Veterinary	15.54	65	Cultural Studies	12.42
31	Materials Engineering	15.50	66	Nursing Science	12.31
32	German Studies	15.30	67	Textile/Clothing Engineering	12.29
33	Forestry	15.20	68	Performing Arts	12.26
34	Mining & Metallurgy	15.18	69	Fine Arts	12.13
35	Political Sciences	15.16	70	Tourism	10.65

Source: German Microcensus 2007-2012; own calculations.

Table A3: Covariate balance treatment and control group, wave 1 (pre-treatment)

Variables	N	All TG	All CG	All Diff. (p-value)	Women TG	Women CG	Women Diff. (p-value)	Men TG	Men CG	Men Diff. (p-value)
Mean (SD) of income of intended field of study	510	17.08 (2.81)	16.97 (2.35)	0.11 (0.76)	16.60 (2.86)	16.54 (2.50)	0.06 (0.88)	17.78 (2.59)	17.64 (1.91)	0.14 (0.77)
Job attribute preference: Very important										
*High income	510	0.32	0.38	-0.06 (0.35)	0.32	0.32	-0.00 (0.94)	0.32	0.46	-0.14 (0.06)
*Contact with others	510	0.22	0.32	-0.10 (0.08)	0.27	0.39	-0.12 (0.02)	0.15	0.21	-0.06 (0.41)
*Helping others	510	0.30	0.28	0.03 (0.60)	0.34	0.33	0.01 (0.86)	0.26	0.20	0.06 (0.51)
*Time for family	510	0.40	0.45	-0.04 (0.41)	0.43	0.45	-0.02 (0.71)	0.36	0.44	-0.07 (0.20)
Educational achievement / skills measures										
Mean (SD) of grade point average	504	3.07 (0.66)	3.02 (0.66)	0.05 (0.62)	3.03 (0.66)	3.01 (0.62)	0.02 (0.84)	3.12 (0.65)	3.04 (0.73)	0.09 (0.54)
Mean (SD) of grade (mathematics)	503	8.53 (3.08)	8.94 (3.25)	-0.41 (0.32)	8.35 (3.11)	8.92 (2.91)	-0.57 (0.22)	8.78 (3.05)	8.96 (3.73)	-0.18 (0.77)
Mean (SD) of grade (German)	499	9.06 (2.44)	9.38 (2.35)	-0.32 (0.40)	9.19 (2.48)	9.61 (2.28)	-0.42 (0.31)	8.88 (2.40)	9.02 (2.41)	-0.15 (0.77)
Mean (SD) of comparative advantage (math vs. German)	499	-0.54 (3.10)	-0.42 (3.37)	-0.12 (0.81)	-0.84 (3.07)	-0.69 (3.27)	-0.15 (0.79)	-0.09 (3.13)	0.00 (3.48)	-0.09 (0.85)
Mean (SD) of figural competence	510	0.03 (1.00)	-0.03 (1.03)	0.06 (0.63)	-0.04 (1.02)	0.00 (1.06)	-0.04 (0.75)	0.12 (0.98)	-0.07 (0.99)	0.20 (0.30)
Mean (SD) of comparative advantage (figural vs. verbal)	510	-0.16 (1.20)	-0.14 (1.18)	-0.02 (0.92)	-0.00 (1.18)	0.04 (1.18)	-0.05 (0.79)	-0.39 (1.19)	-0.43 (1.12)	0.04 (0.86)
Mean (SD) of verbal competence	510	0.19 (1.01)	0.12 (0.96)	0.08 (0.70)	-0.03 (0.97)	-0.04 (0.93)	0.01 (0.97)	0.52 (0.98)	0.36 (0.97)	0.16 (0.47)
Science-oriented course profile	509	0.31	0.10	0.21 (0.14)	0.20	0.06	0.15 (0.17)	0.45	0.16	0.29 (0.14)
Information-related measures										
Feeling informed about HE										
Poorly	510	0.25	0.30	-0.05 (0.32)	0.29	0.33	-0.04 (0.63)	0.20	0.26	-0.07 (0.37)
Partly	510	0.32	0.31	0.01 (0.86)	0.32	0.31	0.01 (0.88)	0.32	0.31	0.00 (0.94)

Well	510	0.43	0.39	0.04	0.39	0.36	0.03	0.48	0.42	0.06
				(0.40)			(0.72)			(0.22)
*Feeling informed about VET										
Poorly	510	0.32	0.36	-0.03	0.29	0.36	-0.07	0.36	0.34	0.02
				(0.58)			(0.29)			(0.79)
Partly	510	0.30	0.36	-0.06	0.34	0.34	-0.00	0.24	0.39	-0.14
				(0.21)			(0.97)			(0.06)
Well	510	0.38	0.29	0.09	0.37	0.30	0.07	0.39	0.27	0.12
				(0.12)			(0.32)			(0.06)
Information sessions in school	510	0.70	0.55	0.15	0.71	0.55	0.15	0.70	0.55	0.14
				(0.20)			(0.23)			(0.34)
*Mean (SD) of number of used sources (private)	510	6.45	6.04	0.42	6.56	6.20	0.36	6.30	5.79	0.51
		(1.87)	(1.96)	(0.07)	(1.58)	(1.75)	(0.12)	(2.24)	(2.22)	(0.13)
Mean (SD) of number of used sources (school)	510	5.22	4.72	0.50	5.21	4.41	0.80	5.23	5.20	0.03
		(2.63)	(2.67)	(0.36)	(2.42)	(2.58)	(0.13)	(2.92)	(2.76)	(0.97)
Further variables										
At least 1 parent with academic degree	510	0.43	0.43	0.00	0.44	0.43	0.01	0.42	0.43	-0.01
				(0.95)	(0.50)	(0.50)	(0.90)	(0.50)	(0.50)	(0.93)
Migration background	507	0.53	0.48	0.04	0.63	0.47	0.16	0.38	0.50	-0.12
				(0.76)			(0.25)			(0.45)
Mean (SD) of risk aversion	509	5.48	5.57	-0.09	5.36	5.37	-0.01	5.65	5.88	-0-23
		(2.19)	(2.21)	(0.62)	(2.11)	(2.12)	(0.97)	(2.31)	(2.31)	(0.46)
*Mean (SD) of external locus of control	510	0.08	-0.06	0.14	0.27	0.03	0.24	-0.18	-0.21	0.02
		(1.10)	(0.98)	(0.25)	(1.09)	(0.94)	(0.10)	(1.05)	(1.01)	(0.88)
School type										
"Gymnasium"	510	0.29	0.37	-0.07	0.36	0.40	-0.04	0.20	0.32	-0.12
				(0.74)			(0.88)			(0.52)
Comprehensive school with a "Gymnasium" track	510	0.38	0.35	0.03	0.42	0.34	0.08	0.32	0.36	-0.05
				(0.90)			(0.74)			(0.83)
Vocational "Gymnasium"	510	0.33	0.28	0.05	0.22	0.26	-0.04	0.48	0.31	0.17
				(0.83)			(0.82)			(0.48)

*Variables used for entropy balancing (substantial part of hypotheses or p-value: <=0.1)

Sources: Best Up, wave 1, analytical sample (see Section 5.3); German Microcensus 2007-2012.